Prefrontal-Parietal-cTBS Effects on Metacognitive Awareness: Real or Illusory?

Antonio Martin[1], Timothy J. Lane[1,2], and Tzu-Yu Hsu[1,2*]

[1] Graduate Institute of Mind, Brain and Consciousness (GIMBC), Taipei Medical University, Taipei, Taiwan

[2] Brain and Consciousness Research Centre (BCRC), TMU-Shuang Ho Hospital, New Taipei City, Taiwan

Corresponding
Tzu-Yu Hsu Ph.D.
Graduate Institute of Mind, Brain and Consciousness, Taipei Medical University
250 Wuxing Street, Taipei, Taiwan.
+886-2-66382736#1233
Email: tzuyu.hsu@tmu.edu.tw

Number of figures:3
Number of tables: 0
Number of words: 3904 (except abstract, figures, tables, and references)

Keywords: visual awareness, metacognition, continuous theta-burst transcranial magnetic stimulation (cTBS), dorsal lateral prefrontal cortex (DLPFC), posterior parietal cortex (PPC).

Abstract

Background: Neuroimaging and lesion studies suggest that the dorsolateral prefrontal and posterior parietal cortices mediate visual metacognitive awareness. The causal evidence provided by non-invasive brain stimulation, however, is inconsistent.

Objective/Hypothesis: Here we revisit major experiments with an improved methodology and a new Kanizsa figure task trying to determine whether bilateral continuous theta-burst transcranial magnetic stimulation (cTBS) over these regions affects perceptual metacognition. Specifically, we tested whether subjective visibility ratings and/or metacognitive efficiency are lower when cTBS is applied to these two regions in comparison to an active control region.

Methods: A within-subjects design including three sessions, at one-week intervals, was implemented. In each session, every participant was administered bilateral cTBS to either prefrontal, control or parietal cortices. Two concurrent tasks were performed: real and illusory figure, using a staircase procedure to stabilize objective performance.

Results: We found strong null findings for the replicated task, suggesting that cTBS affected neither visibility ratings nor visual metacognition. Nevertheless, we also observed that the implementation of illusory Kanizsa figures disrupted metacognitive ability, specifically when cTBS was applied to prefrontal but not parietal areas.

Conclusion(s): This study adds new evidence to the debate regarding the role of the PPN in metacognition: specifically, using rigorous criteria, evidence is adduced to show that the DLPFC cTBS selectively interferes with metacognitive awareness. What is more, a novel method for investigating the role of the PPN in metacognition is presented.

Introduction

Visual awareness of perceptual contents has been suggested to require involvement of first person monitoring of one's own mental processing (Brown et al., 2019). One approach to evaluating visual awareness is use of metacognitive judgements (i.e., subjective ratings), including visibility or confidence, along with objective performance. Such an approach can be implemented in various cognitive experiments, including visual discrimination tasks, which allow us to investigate metacognitive sensitivity, or how well metacognitive judgements can predict objective performance. The neural correlates of metacognitive judgements and metacognitive sensitivity both seem to involve the prefrontal parietal network, PPN (Boly et al., 2017; Bor & Seth, 2012; Odegaard et al., 2017). The PPN also is sometimes referred to as the central executive network (Ryali et al., 2016) or the lateral frontoparietal network (Uddin et al., 2019); no matter which term is used, the relevant regions include the dorsolateral prefrontal and the posterior parietal cortices (DLPFC and PPC). But causal studies—as implemented by non-invasive brain stimulation (NIBS)—suggest that the DLPFC mediates metacognitive judgements (Shekhar & Rahnev, 2018) or metacognitive sensitivity (Rounis et al., 2010; Ruby et al., 2018). But these findings have been challenged (Bor et al., 2017, 2018). To our knowledge, this challenge has not been addressed adequately yet.

A first line of evidence derives from correlational and lesion studies that show patients with a damaged prefrontal cortex who have intact objective performance but selective impairment of metacognitive awareness (Colás et al., 2019; Del Cul et al., 2009; Grossner et al., 2018). For example, Colás et al. (2019) observed that patients with DLPFC damage, despite having normal levels of objective spatial orienting abilities, presented a more conservative response bias affecting their subjective visibility threshold. Consistent with these lesion studies, prefrontal cortex grey matter volume and neural activity in healthy participants also evinces positive correlations with confidence ratings or metacognitive sensitivity, but not with objective performance (Fleming et al., 2010, 2012). Furthermore, parietal cortex neural activity, especially the PPC, is also associated with confidence level (Hanks et al., 2011; Kiani & Shadlen, 2009; Pereira et al., 2020; Simons et al., 2010). Simons et al. (2010) observed specific decreases in confidence levels after bilateral impairments to the PPC, similar to those described for the DLPFC damages. These results converge in suggesting that the PPN is essential to metacognition.

A second line of evidence derived from NIBS studies lends further support to the claim that the DLPFC plays a causal role in metacognitive awareness. Rounis and her colleagues (2010) applied bilateral offline continuous theta-burst transcranial magnetic stimulation (cTBS) to temporarily disrupt PPN activity, specifically in the DLPFC, in order to evaluate the DLPFC's effect on metacognitive sensitivity. They observed that when controlling for objective performance, bilateral DLPFC cTBS affected visibility ratings. This dissociation between performance and visibility is an instance of impaired metacognitive sensitivity. The authors hypothesized that the DLPFC is causally involved in metacognitive sensitivity. An opposing view derives from Shekhar and Rahnev (2018), who used online transcranial magnetic stimulation (TMS) and concluded that the anterior PFC—but not the DLPFC—plays a pivotal role in metacognitive sensitivity. They did find, however, that the DLPFC modulates all confidence ratings. Based upon this finding they hypothesize that the DLPFC causally affects the strength of the sensory evidence. Strikingly, Bor et al. (2017), in both between- and within-subject studies, failed to replicate any of these findings, suggesting that neither the DLPFC nor the PPC are playing the conjectured roles.

Several methodological differences might explain the failure to replicate the effects of cTBS in metacognition, particularly the exclusion of unstable datasets (Ruby et al., 2018; Bor et al., 2018). In this investigation, we attempted to refine experimental methodology and question whether the previously used tasks are suited to investigating the effects of cTBS over metacognitive awareness. For this purpose, we decided to compare the original task to a similar one, but one that employs Kanizsa illusory contours, so as to potentially elevate the effects of cTBS over metacognition. The DLPFC has been suggested to contribute to the formation of expectations that shape subsequent responses, particularly encoding the reliability of expected sensory inputs (de Lange et al., 2018; Lobanov et al., 2014; Tanaka et al., 2006). According to the predictive coding hypothesis (Friston & Kiebel, 2009), when bottom-up sensory inputs are ambiguous or unreliable—as in the case of illusory contours—information from internal models tends to be weighted more heavily (Parr et al., 2018; Yu, 2014). It follows from this hypothesis that altering the DLPFC activation by means of cTBS should disrupt metacognition when stimuli are ambiguous (Sherman et al., 2015).

In sum, to date NIBS investigations of whether cTBS applied to the PPN negatively affects metacognitive awareness have yielded contradictory results. To make progress on resolving

these contradictions, for this study, we examine how the PPN is related to metacognitive visual awareness. In order to do so, we replicated the Rounis et al. (2010) figure discrimination task and compared it to an illusory Kanizsa figure task (Kanizsa, 1976). When three or four 'pacmen'-like inducers are arranged in a configuration, they can create the illusion that a shape exists, even absent an actual physical shape. We chose Kanizsa style figures because of their similarity to the original task and because they are assumed to depend on top-down contextual neural activity associated with unreliable bottom-up sensory inputs (Banica & Schwarzkopf, 2016; Kok et al., 2016). Adhering to a within-subjects design, offline cTBS was administered bilaterally at either the DLPFC or the PPC, during different sessions, and then compared to an active control site, the somatosensory cortex (S1).

Method

Participants

Although a total of thirty-five participants were recruited for this study, all of whom met our inclusion criteria (see Supplementary information S2), two of them could not finish the experiment due to the COVID-19 epidemic. Although thirty-three participants successfully completed the experiment, twelve participants were excluded due to low or unstable task performance (see Supplementary information S3). Our final list comprised 11 females and 10 males, ages ranging from 23 to 41, averaging 30.92 (SD = 5.23). Age difference between genders was not statistically significant ($t_{17.61} = 1.41$, $p = 0.18$). All participants had normal or corrected-to-normal vision. Written consent was obtained from all participants prior to their participation in this study. All participants were informed about the purpose of the study and procedures before being asked to give consent. This study was approved by the joint institutional Review Board, Taipei Medical University, Taipei City, Taiwan (N201910050).

Apparatus

Stimuli were generated and delivered by PsychoPy2 Experiment Coder v1.90.3 (Peirce et al., 2019), running under Python, on an LCD monitor with a vertical refresh rate of 60 Hz. Participants viewed the stimuli, head position stabilized by a chin-rest, at a distance of 40 cm. All statistical analyses were performed with R 4.0.3 and JASP 0.16.3.

Stimuli and visual discrimination tasks

The Rounis et al. (2010) experimental design was adopted, with minor adjustments following Bor et al. (2017), a design described in the Procedure section. Participants were required to

perform a two-alternative, forced-choice visual discrimination task (see Figure 1A) in a dark room. For the real figure task, two 0.8^{o} visual angle black stimuli, a diamond and a square, were presented simultaneously, on a grey background, 33 ms for both left and right, 1^{0} visual angle away from the central fixation cross, followed by a 50 ms mask. Participants were asked to discriminate the location of a square or a diamond, either on the left or right, while fixating their gaze on a fixation cross in the centre of the screen. For the illusory figure task, the stimuli were replaced by 0.425^{o} radius wide 'pacmen'-like inducers at each vertex to induce 0.8^{o} wide square-like or diamond-like Kanizsa illusory shapes for 83 ms (see Figure 1B). This too was followed by a 50 ms mask. Giving similar instructions for both tasks, participants were asked to determine whether the square was on the left and the diamond on the right (left keys: 'c' or 'v'), or vice versa (right keys: 'n' or 'm'), while fixating their gaze on the fixation cross in the centre of the screen. The location of these opposing figures was pseudo randomized with equal probability. Simultaneously, participants were asked to judge the visibility of the stimuli. If the stimuli in the current trial were felt to be more visible, relative to others in this experiment, the participant was asked to report "clear" by pressing keys: 'c' (if the target appeared on the left visual field) or 'm' (if the target appeared on the right visual field). Alternatively, they were asked to press the "unclear" button 'v' (if the target appeared in the left visual field) or 'n' (if the target appeared in the right visual field).

Although several parameters have been developed to assess metacognitive sensitivity (Fleming & Lau, 2014), we followed the meta-d' parameter estimation, since it has been argued to be unaffected by response bias and to provide a more transparent measure of metacognitive ability (Maniscalco & Lau, 2012). In order to calculate the meta-d' values, visibility reports were taken to generate approximations of what the type1 d' values would be, when using the sum of square errors approximation (SSE method).

Theta-burst stimulation

The cTBS pulses were administered with a Magpro X100 (MagVenture, Denmark) in a 70 mm figure-of-eight-shaped coil (MC-B70, MagVenture). Each bilateral cTBS session comprises 600 pulses, at 80% intensity stimulation for adjusted active motor threshold (adjAMT), and were applied over the left and right DLPFC, PPC, or S1.

Every experimental session comprises the administration of cTBS to one region bilaterally. The sequence of stimulation was counterbalanced for each participant across the three

experimental sessions. The MNI coordinates for bilateral DLPFC were right [36 30 50] and left [-42 20 50] (Cieslik et al., 2013; Scheperjans et al., 2008); for bilateral PPC, right [36 -53 50] and left [-34 -52 50]; and, for the S1 (control site), right [22 -32 70] and left [-22 -30 70] (see Figure 2). An average observed adjAMT of 28.92 (SD = 5.81) led to cTBS intensities of 24.44 (SD = 7.32) for the DLPFC, 41.96 (SD = 9.07) for the PPC, and 47.83 (SD = 9.91) for the S1. More information regarding how the adjAMT was calculated for each region, depending on its depth, can be found in Supplementary information (S1).

cTBS delivers a 20 s train of uninterrupted theta-bursts pulses. This consists of 3 pulses, at 50 Hz, given at 200 ms intervals, comprising a total of 300 pulses for 20 seconds (or 100 bursts), applied to one hemisphere. After a one- minute inter-stimulation interval, the procedure was repeated in the corresponding location of the other hemisphere. The coil was attached, perpendicular to the location of stimulation (90^0 angle), and the task was initiated about 1 minute after cTBS administration.

Magnetic resonance imaging

Before participants began the three main experimental sessions, anatomical images for each were collected, using a 3T magnetic resonance scanner, either Siemens MAGNETOM Skyra 3T or GE 3.0 Tesla Discovery MR750, depending on accessibility. Their T1-weighted images were used to locate the regions for administration of the cTBS as well as for adjAMT calculations.

Procedure

Each participant was required to come to the laboratory on four different days. On the first day, participants were introduced to the rationale of the experiment and allowed to practice, to ensure their understanding of the study. After providing informed consent, both their anatomical images and their AMT were collected. Over the subsequent three days, participants were required to do as many practice blocks as necessary, each comprising 100 trials, in order to reach a steady performance state. Their performance on the discrimination tasks was set to 75% manipulating the contrast of the stimuli, using a staircase procedure. For each trial, the target contrast increases whenever an incorrect response is made, and decreases after two or three consecutive correct responses. The staircase adaptations were larger for the initial practice trials (step size at 25%), followed by a constant step size of 5%, used in the experiment blocks. Participants were told that the location of the figures was randomised.

Prior to cTBS administration, many practice blocks were used, to ensure that participants had achieved a steady performance of 75% accuracy. After cTBS administration, participants were required to perform three blocks of real/illusory task, followed by three more of the other task. Each block comprises 100 trials. A short break of one minute follows at the end of each block. The order of the tasks and brain sites was counterbalanced across participants and sessions (i.e. cTBS sites).

Statistical Analyses

In order to determine whether cTBS negatively impacts metacognitive awareness, directional and planned comparisons are implemented for both overall visibility ratings and metacognitive efficiency values, while taking S1-cTBS as a control condition. These comparisons assess whether S1-cTBS visibility ratings or metacognitive efficiencies are higher than DLPFC or PPC-cTBS. All statistical analyses follow a frequentist repeated measures ANOVA, followed by a Bayesian calculation of the Bayes Factor (BF_{10}), representing the relative strength of evidence for the alternative with respect to the null hypothesis. For other statistical analyses (such as accuracy or response bias) no particular directionality is assessed and thus all statistical analyses are ad hoc, with Bonferroni corrections for Type1 errors. Statistical analyses were corrected, if parametrical assumptions of centrality or sphericity were violated, by using either the nonparametrical alternative or the Greenhouse-Geisser correction.

Results

Overall levels of accuracy were not statistically different, neither between real and illusory figure tasks ($F(1, 20) = 0.55$, $p = 0.47$, $BF_{10} = 0.42$), nor between cTBS administration sites ($F(2, 40) = 0.09$, $p = 0.91$, $BF_{10} = 0.14$; see Figure 3A). The mean contrast level to keep accuracy at 75% did not change across cTBS sites ($F(2, 40) = 0.01$, $p = 0.99$, $BF_{10} = 0.13$), nor between illusory and real figure tasks ($F(1, 20) = 0.99$, $p = 0.33$, $BF_{10} = 0.47$). Similarly, reaction times (RTs) were not observed to differ across cTBS sites ($F(2, 40) = 0.22$, $p = 0.80$, $BF_{10} = 0.15$), between task types ($F(1, 20) = 0.06$, $p = 0.81$, $BF_{10} = 0.30$). The main effect of blocks was significant ($F(1.45, 29.08) = 13.07$, $p < 0.001$, $\eta_p^2 = 0.40$, $BF_{10} = > 99$). The polynomial contrast showed a linear decrement in RTs with increasing number of blocks ($t(20) = 4.86$, $p < 0.001$, *Cohen's d* = 0.24, $BF_{10} = 34.54$). Discrimination task sensitivities (d'), with a mean of 1.43 (SD = 0.18), were not affected by cTBS administration ($F(2, 40) = 0.05$, $p = 0.95$, $BF_{10} = 0.14$) and did not differ between illusory and real figure tasks ($F(1, 20)$

= 1.89, p = 0.18, BF_{10} = 0.71; see Figure 3B). Similar negative results were also found for the unbiased response criterion (ca; $F(2, 40) = 0.80$, $p = 0.46$, $BF_{10} = 0.23$, and $F(1, 20) < 0.01$, $p = 0.97$, $BF_{10} = 0.31$ respectively), with a mean of -0.02 (SD = 0.17).

Metacognitive ratings of visibility were significantly lower for incorrect trials ($t(20) = 5.44$, $p < 0.001$, *Cohen's d* = 1.19, $BF_{10} > 99$; see Figure 3C), with no observed differences between illusory and real figure tasks ($F(1, 20) = 0.05$, $p = 0.83$, $BF_{10} = 0.31$). The mean phi correlations between discrimination accuracy and visibility ratings did not change: neither between cTBS sites ($F(2, 40) = 0.05$, $p = 0.19$, $BF_{10} = 0.13$), nor between tasks ($F(1, 20) = 1.87$, $p = 0.19$, $BF_{10} = 0.60$), averaging 0.14 (SD = 0.11). The overall ratings of visibility did not change: neither due to the nature of the task being performed ($F(1, 20) = 0.05$, $p = 0.83$, $BF_{10} = 0.31$), nor due to the cTBS administration site ($F(2, 40) = 0.77$, $p = 0.47$, $BF_{10} = 0.23$). Even though there was an increase in visibility ratings specific to the illusory figure task, the conjectured contrast among cTBS sites (S1 vs DLPFC, PPC) was not significant ($t(40) = 1.17$, $p = 0.25$, $BF_{10} = 0.29$), thereby suggesting that cTBS did not affect reports of visibility.

Metacognitive efficiency values were lower for the illusory figure task in comparison to the real figure task ($F(1, 20) = 4.86$, $p < 0.05$, $\eta_p^2 = 0.20$, $BF_{10} = 1.73$), with averages of –0.82 (SD = 0.56) and –0.60 (SD = 0.64). The order of tasks did not affect metacognitive efficiency ($F(1, 20) = 0.05$, $p = 0.82$, $BF_{10} = 0.31$). For the real figure task, no significant effects of cTBS were observed on metacognitive efficiency ($F(2, 40) = 0.92$, $p = 0.41$, $BF_{10} = 0.25$; see Figure 3D). For the illusory figure task, the Grubb's test identified a participant presenting a significantly deviant value (G = 2.67, $p < 0.05$) which also affected the normality assumption (W = 0.93, $p < 0.05$). After the removal of this outlier, it was observed that cTBS administered over the DLPFC significantly reduced metacognitive efficiencies when compared to the S1 ($t(19) = 1.89$, $p < 0.05$, *Cohen's d* = 0.42, $BF_{10} = 1.96$). Administration of cTBS to the PPC did not change metacognitive efficiencies when compared to the S1 ($t(19) = 0.01$, $p = 0.51$, $BF_{10} = 0.23$).

Discussion

This study was designed to investigate whether cTBS over the DLPFC or the PPC could diminish either visibility ratings or metacognitive sensitivities, when compared to an active control region (S1) and while sparing objective task performance. Our results indicated that

objective task accuracy, discrimination sensitivity and task difficulty were not affected by cTBS administration, neither for real nor for illusory figure tasks. Subjective visibility reports were not affected by cTBS administration site, neither for real nor for illusory figure tasks. Interestingly, cTBS over the DLPFC, but not over the PPC, significantly reduced metacognitive efficiencies only when performing the illusory task, and not when performing the real figure task.

cTBS administration did not affect visibility ratings, supporting the idea that visibility reports may not be influenced by the integrity of PPNs, at least for this real figure task. These null results replicated previous findings with confidence ratings using a similar task design (Bor et al., 2017, 2018). Although a reduction in confidence ratings has been reported when applying online TMS to the DLPFC (Shekhar and Rahnev, 2018), unlike our study, this finding was based on a task that lacked a staircase procedure for continuous stabilization of objective performance. Moreover, that study had a broader time window between stimulus presentation and response; this greater lag time could have allowed post-perceptual cognitive processes (i.e., a read-out function) to create the reported effects.

In this study, we attempted to replicate the previous study proposed by Rounis et al (2010) by introducing improvements and addressing several design concerns raised by Bor et al. (2017). Our task adopted the best of both designs: as suggested by Bor et al. (2010), participants practiced, using as many blocks as necessary, in order to achieve a steady performance, thereby avoiding unstable datasets. What is more, following Rounis et al. (2010), we adhered to 5% increments in our adaptive staircase in order to allow for quick task difficulty adjustment. Accordingly, it was only necessary to exclude three participants due to unstable datasets, whereas most exclusion was due to an inappropriate accuracy level (around 50%). Despite the fact that the data exclusion criteria result in the loss of many participants, these criteria prior a NIB study aid in ensuring the quality of the data.

Even though all measures of objective task performance for the two tasks were equivalent, Kanizsa illusory contours showed a reduction in metacognitive efficiency. To the best of our knowledge, this is the first time that Kanizsa illusory contours have been compared to a real figure task such that an effect on metacognition is exhibited. Strictly speaking, relative to what we consciously experience, all representations in the world are incomplete, albeit some more than others, textbook illusions being those that comprise a relatively greater degree of

incomplete stimuli. Conscious “filling in” of the gaps for these designed illusory figures may crucially depend on top-down priors generated at higher cortical regions, such as the lateral occipital cortex (Murray & Herrmann, 2013). We hypothesize that this effect is due to a reduction in the bottom-up reliability of illusory stimuli, which is in line with previous findings on the role that expectations have over metacognition (Sherman et al., 2015). More work is needed, however, in order to confirm the claim that illusory percepts are in fact less reliable than are other figures.

The DLPFC has been found to mediate many functions, including attention (Buschman & Miller, 2009; Brandt et al., 1998; Corbetta & Shulman, 2002), working memory (Miller, Lundqvist, and Bastos, 2018), and the construction of contextual expectations or priors from a top-down perspective (Rahnev et al., 2011; Banica & Schwarzkopf, 2016; Kok et al., 2016). The present study is the first to demonstrate that the DLPFC cTBS selectively interferes with metacognitive efficiency when rigorous exclusion criteria are adopted. We replicated the effect that Rounis et al. (2010) reported with illusory figures. This is consistent with the notion that the prefrontal cortex is involved in metacognition (Brown et al., 2019; Rahnev, 2017). Additionally, internal models play a significant role in creating an illusory contour (Parr et al., 2018; Yu, 2014) as well as in determining an individual's objective performance, both when external stimuli provide only incomplete information and when external feedback is unavailable. Although it is unclear which internal models the DLPFC cTBS interferes with in the current design, it does seem that manipulation with limited external stimuli or no external feedback might serve as an effective NIBS experimental design. In sum, this study adds new evidence to the debate regarding the role of the PPN in metacognitive awareness (Rounis et al, 2010; Bor et al., 2017; Boly et al., 2017; Odegaard et al., 2017), and it presents a new approach to addressing the role of the PPN in metacognition.

References

Banica, T., & Schwarzkopf, D. S. (2016). Induction of Kanizsa Contours Requires Awareness of the Inducing Context. PLOS ONE, 11(8), e0161177. https://doi.org/10.1371/journal.pone.0161177

Barch, D. M., Braver, T. S., Nystrom, L. E., Forman, S. D., Noll, D. C., & Cohen, J. D. (1997). Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia*, *35*(10), 1373–1380. https://doi.org/10.1016/S0028-3932(97)00072-9

Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., & Tononi, G. (2017). Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. Journal of Neuroscience, 37(40), 9603–9613. https://doi.org/10.1523/JNEUROSCI.3218-16.2017

Bor, D., Barrett, A. B., Schwartzman, D. J., & Seth, A. K. (2018). Response to Ruby et al: On a 'failed' attempt to manipulate conscious perception with transcranial magnetic stimulation to prefrontal cortex. Consciousness and Cognition, 65, 334–341. https://doi.org/10.1016/J.CONCOG.2018.07.011

Bor, D., Schwartzman, D. J., Barrett, A. B., & Seth, A. K. (2017). Theta-burst transcranial magnetic stimulation to the prefrontal or parietal cortex does not impair metacognitive visual awareness. PLOS ONE, 12(2), e0171793. https://doi.org/10.1371/journal.pone.0171793

Bor, D., & Seth, A. K. (2012). Consciousness and the Prefrontal Parietal Network: Insights from Attention, Working Memory, and Chunking. Frontiers in Psychology, 3, 63. https://doi.org/10.3389/fpsyg.2012.00063

Brown, R., Lau, H., & LeDoux, J. E. (2019). Understanding the Higher-Order Approach to Consciousness. Trends in Cognitive Sciences, 23(9), 754–768. https://doi.org/10.1016/J.TICS.2019.06.009

Brandt, S. A., Ploner, C. J., Meyer, B.-U., Leistner, S., & Villringer, A. (1998). Effects of repetitive transcranial magnetic stimulation over dorsolateral prefrontal and posterior parietal cortex on memory-guided saccades. *Experimental Brain Research*, *118*(2), 197–204. https://doi.org/10.1007/s002210050272

Buschman, T. J., & Miller, E. K. (2009). Serial, Covert Shifts of Attention during Visual Search Are Reflected by the Frontal Eye Fields and Correlated with Population Oscillations. *Neuron*, *63*(3), 386–396. https://doi.org/10.1016/j.neuron.2009.06.020

Cieslik, E. C., Zilles, K., Caspers, S., Roski, C., Kellermann, T. S., Jakobs, O., ... Eickhoff, S. B. (2013). Is there “one”; DLPFC in cognitive action control? Evidence for heterogeneity from co-activation-based parcellation. Cerebral Cortex (New York, N.Y. : 1991), 23(11), 2677–2689. https://doi.org/10.1093/ cercor/bhs256

Colás, I., Chica, A. B., Ródenas, E., Busquier, H., Olivares, G., & Triviño, M. (2019). Conscious perception in patients with prefrontal damage. Neuropsychologia, 129, 284–293. https://doi.org/10.1016/j.neuropsychologia.2019.03.002

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

De Lange, F. P., Heilbron, M., & Kok, P. (2018, September 1). How Do Expectations Shape Perception? Trends in Cognitive Sciences, Vol. 22, pp. 764–779. https://doi.org/10.1016/j.tics.2018.06.002

Del Cul, A., Dehaene, S., Reyes, P., Bravo, E., & Slachevsky, A. (2009). Causal role of prefrontal cortex in the threshold for access to consciousness. Brain, 132(9), 2531–2540. https://doi.org/10.1093/brain/awp111

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39(2), 175–191. https://doi.org/10.3758/BF03193146

Fleming, S. M., Huijgen, J., & Dolan, R. J. (2012). Prefrontal contributions to metacognition in perceptual decision making. Journal of Neuroscience, 32(18), 6117–6125. https://doi.org/10.1523/JNEUROSCI.6489-11.2012

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. Frontiers in Human Neuroscience, 8, 443. https://doi.org/10.3389/fnhum.2014.00443

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. Science, 329(5998), 1541–1543. https://doi.org/10.1126/science.1191883

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1521), 1211–1221. https://doi.org/10.1098/rstb.2008.0300

Grossner, E. C., Bernier, R. A., Brenner, E. K., Chiou, K. S., & Hillary, F. G. (2018). Prefrontal gray matter volume predicts metacognitive accuracy following traumatic brain injury. Neuropsychology, 32(4), 484–494. https://doi.org/10.1037/neu0000446

Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E., & Shadlen, M. N. (2011). Elapsed decision time affects the weighting of prior probability in a perceptual decision task. Journal of Neuroscience, 31(17), 6339–6352. https://doi.org/10.1523/JNEUROSCI.5613-10.2011

Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. Perception & Psychophysics, 49(3), 227–229. https://doi.org/10.3758/BF03214307

Kanizsa, G. (1976). Subjective Contours. Scientific American, Vol. 234, pp. 48– 53. https://doi.org/10.2307/24950327

Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. Science (New York, N.Y.), 324(5928), 759–764. https://doi.org/10.1126/science.1169405

Kok, P., Bains, L. J., van Mourik, T., Norris, D. G., & de Lange, F. P. (2016). Selective Activation of the Deep Layers of the Human Primary Visual Cortex by Top-Down Feedback. Current Biology, 26(3), 371–376. https://doi.org/10.1016/J.CUB.2015.12.038

Lobanov, O. V., Zeidan, F., McHaffie, J. G., Kraft, R. A., & Coghill, R. C. (2014). From cue to meaning: Brain mechanisms supporting the construction of expectations of pain. Pain, 155(1), 129–136. https://doi.org/10.1016/j.pain.2013.09.014

Lowe, C. J., Manocchio, F., Safati, A. B., & Hall, P. A. (2018, March 1). The effects of theta burst stimulation (TBS) targeting the prefrontal cortex on executive functioning: A systematic review and meta-analysis. Neuropsychologia, Vol. 111, pp. 344–359. https://doi.org/10.1016/j.neuropsychologia.2018.02.004

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. Consciousness and Cognition, 21(1), 422–430. https://doi.org/10.1016/J.CONCOG.2011.09.021

Miller EK, Lundqvist M, Bastos AM. (2018). Working Memory 2.0. Neuron. 100(2), 463-475. doi: 10.1016/j.neuron.2018.09.023.

Murray, M. M., & Herrmann, C. S. (2013). Illusory contours: a window onto the neurophysiology of constructing perception. *Trends in Cognitive Sciences*, *17*(9), 471–481. http://dx.doi.org/10.1016/j.tics.2013.07.004

Odegaard, B., Knight, R. T., & Lau, H. (2017). Should a Few Null Findings Falsify Prefrontal Theories of Conscious Perception? The Journal of Neuroscience. https://doi.org/10.1523/JNEUROSCI.3217-16.2017

Parr, T., Benrimoh, D. A., Vincent, P., & Friston, K. J. (2018). Precision and False Perceptual Inference. Frontiers in Integrative Neuroscience, 12: 39. doi: 10.3389/fnint.2018.00039

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. Behavior Research Methods, 51(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pereira, M., Megevand, P., Tan, M. X., Chang, W., Wang, S., Rezai, A., ... Faivre, N. (2020). Evidence accumulation determines conscious access. BioRxiv. https://doi.org/10.1101/2020.07.10.196659

Rahnev, D., Evan, D. N., Riddle, J., Sue, A. L., D'Esposito, M. (2017). Causal evidence for frontal cortex organization for perceptual decision making. *Proceedings of the National Academy of Sciences*, *113*(21), 6059–6064. https://doi.org/10.1073/pnas.1522551113

Rahnev, D., Maniscalco, B., Graves, T., Huang, E., de Lange, F. P., Lau, H. (2011). Attention induces conservative subjective biases in visual perception. *Nature Neuroscience*, *14*, 1513–1515. https://doi.org/10.1038/nn.2948

Rosero Pahi, M., Cavalli, J., Nees, F., Flor, H., & Andoh, J. (2020). Disruption of the Prefrontal Cortex Improves Implicit Contextual Memory-Guided Attention: Combined Behavioral and Electrophysiological Evidence. *Cerebral Cortex*, *30*(1), 20–30. https://doi.org/10.1093/cercor/bhz067

Rossi, S., Hallett, M., Rossini, P. M., & Pascual-Leone, A. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. Clinical Neurophysiology, 120(12), 2008–2039. https://doi.org/10.1016/J.CLINPH.2009.08.016

Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. Cognitive Neuroscience, 1(3), 165–175. https://doi.org/10.1080/17588921003632529

Ruby, E., Maniscalco, B., & Peters, M. A. K. (2018). On a ‘failed’ attempt to manipulate visual metacognition with transcranial magnetic stimulation to prefrontal cortex. Consciousness and Cognition, 62, 34–41. https://doi.org/10.1016/J.CONCOG.2018.04.009

Ryali, S., Supekar, K., Chen, T., Kochalka, J., Cai, W., Nicholas, J., ... Menon, V. (2016). Temporal Dynamics and Developmental Maturation of Salience, Default and Central-Executive Network Interactions Revealed by Variational Bayes Hidden Markov

Modeling. PLOS Computational Biology, 12(12), e1005138. https://doi.org/10.1371/journal.pcbi.1005138

Scheperjans, F., Eickhoff, S. B., Homke, L., Mohlberg, H., Hermann, K., Amunts, K., & Zilles, K. (2008). Probabilistic Maps, Morphometry, and Variability of Cytoarchitectonic Areas in the Human Superior Parietal Cortex. Cerebral Cortex, 18(9), 2141–2157. https://doi.org/10.1093/cercor/bhm241

Shekhar, M., & Rahnev, D. (2018). Distinguishing the Roles of Dorsolateral and Anterior PFC in Visual Metacognition. The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 38(22), 5078–5087. https:// doi.org/10.1523/JNEUROSCI.3484-17.2018

Sherman, M. T., Seth, A. K., Barrett, A. B., & Kanai, R. (2015). Prior expectations facilitate metacognition for perceptual decisions. Consciousness and Cognition, 35: 53–65.

Simons, J. S., Peers, P. V., Mazuz, Y. S., Berryhill, M. E., & Olson, I. R. (2010). Dissociation between memory accuracy and memory confidence following bilateral parietal lesions. Cerebral Cortex, 20(2), 479–485. https://doi.org/10.1093/cercor/bhp116

Stokes, M. G., Chambers, C. D., Gould, I. C., English, T., McNaught, E., McDonald, O., & Mattingley, J. B. (2007). Distance-adjusted motor threshold for transcranial magnetic stimulation. Clinical Neurophysiology, 118(7), 1617–1625. https://doi.org/10.1016/j.clinph.2007.04.004

Stokes, M. G., Chambers, C. D., Gould, I. C., Henderson, T. R., Janko, N. E., Allen, N. B., & Mattingley, J. B. (2005). Simple metric for scaling motor threshold based on scalp-cortex distance: Application to studies using transcranial magnetic stimulation. Journal of Neurophysiology, 94(6), 4520– 4527. https://doi.org/10.1152/jn.00067.2005

Tanaka, S. C., Samejima, K., Okada, G., Ueda, K., Okamoto, Y., Yamawaki, S., & Doya, K. (2006). Brain mechanism of reward prediction under predictable and unpredictable environmental dynamics. Neural Networks, 19(8), 1233– 1241. https://doi.org/10.1016/j.neunet.2006.05.039

Uddin, L. Q., Yeo, B. T. T., & Spreng, R. N. (2019). Towards a Universal Taxonomy of Macro-scale Functional Human Brain Networks. Brain Topography, Vol. 32, pp. 926–942. https://doi.org/10.1007/s10548- 019-00744-6

Yu, A. J. (2014). Bayesian models of attention. In A. C. Nobre & S. Kastner (Eds.), Oxford library of psychology. The Oxford handbook of attention (p. 1159–1197). Oxford University Press.

A B

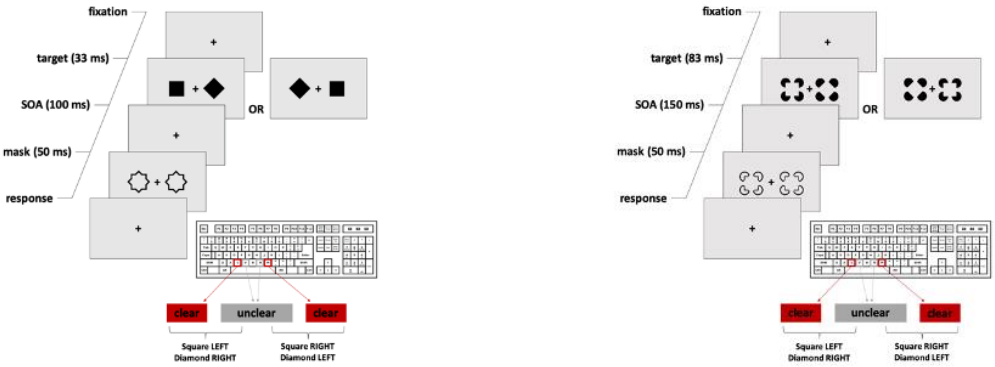

Figure 1 Experimental Design. (A) Replicated two-alternative forced-choice visual task (from Rounis et al., 2010). Participants were asked to discriminate the location of the square to be in either left or right to the fixation cross. Simultaneously, participants also offer a rating over how strong is their subjective visibility ("clear" or "unclear"). (B) New version of the visual task. In this second experiment, participants were given the same instructions but need to locate the illusory square.

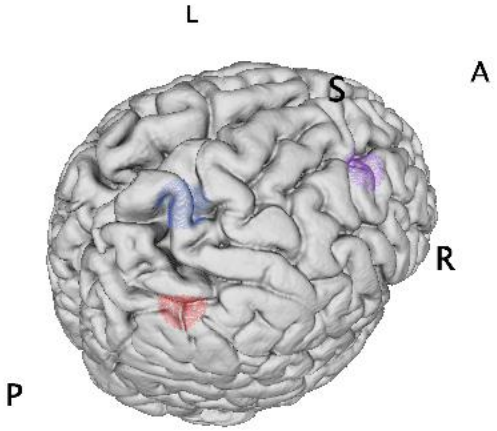

Figure 2. cTBS Stimulation Sites. Right hemisphere view of bilaterally targeted sites: DLPFC (purple), PPC (red), and S1 (blue).

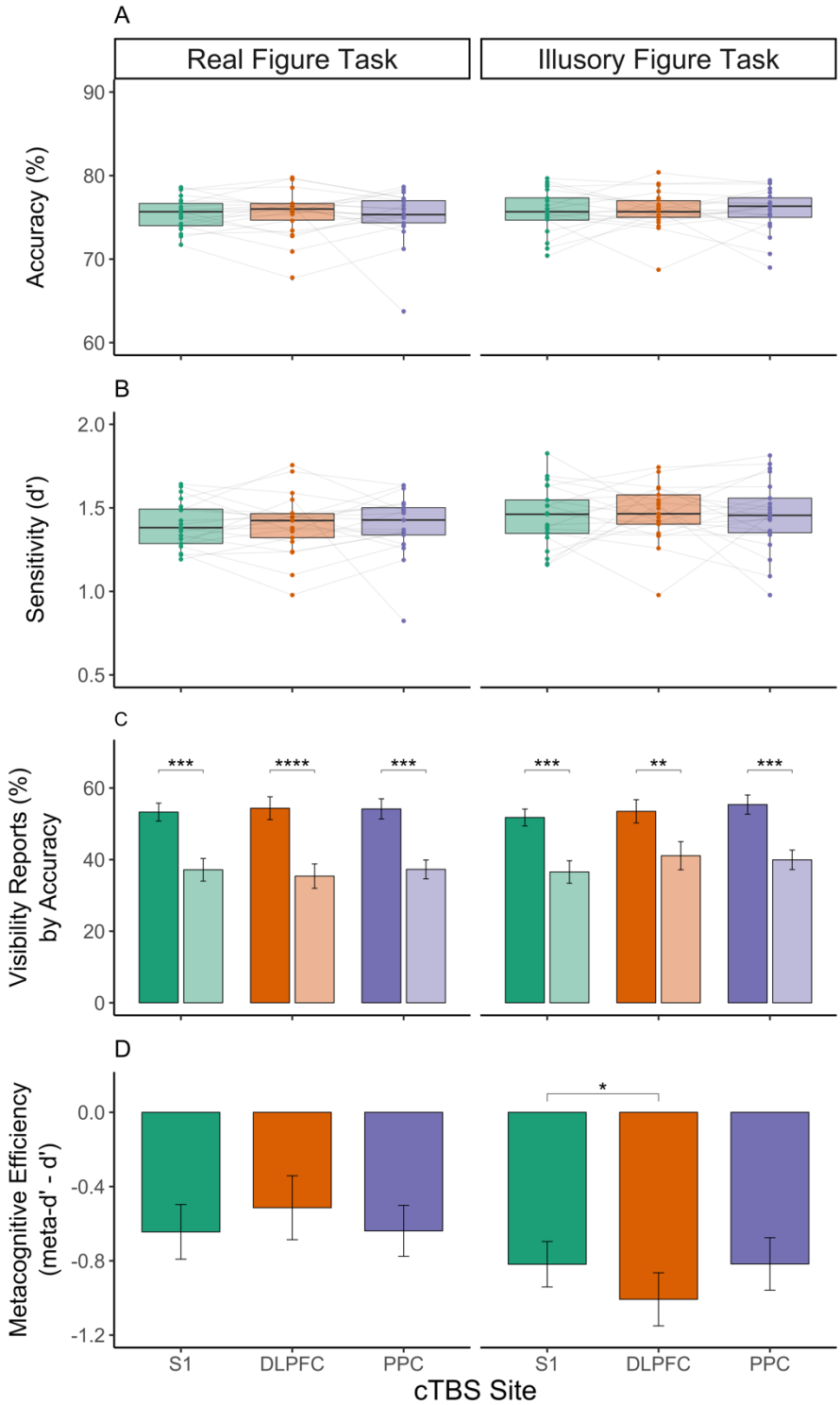

Figure 3. (A) Discrimination task accuracy was successfully stabilised with the use of an adaptive staircase. (B) Which also stabilised perceptual sensitivity (d'). (C) Visibility reports, in turn, were allowed to vary and were observed to be particularly higher at correct (plane bars) compared to incorrect trials (opaque bars). (D) Interestingly, cTBS administration significantly reduced metacognitive abilities when applied to the DLPFC only at the illusory figure task. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$

Acknowledgments

Special thanks to Dr. Daniel Bor and Prof. Hakwan Lau for sharing their Matlab scripts for the real figure discrimination task which was rewritten with python language, as well as to Dr. Niall W. Duncan for his help building the scalpdist software with FSL analyses tool for MRI which allowed us to estimate the distance from a cortical region in MNI coordinates to the scalp. Also, we would like to thank Ko-Ping Chou, Hsin-Yi Wang, and Paul Z. Cheng for their assistance in collecting data.
This work is sponsored by grants from the National Science and Technology Council, Taiwan (111-2410-H-038-012-3) to Professor Timothy Lane and (109-2410-H-038-010; 111-2410-H-038-009-MY2) to Dr. Tzu-Yu Hsu.

Author contributions and competing interests' declaration

AM and TYH designed the study. AM coded all the scripts, collected and analysed the data. AM, TY and TL wrote, revised and approved the submitted manuscript. The authors declare no conflicts of interest.

Supplementary information

S1: Calculation of the adjusted Active Motor Threshold (adjAMT)

The motor evoked potential (MEP) was elicited by placing the coil oriented 45 degrees to the coronal plane and measured from the right first dorsal interosseous (FDI) hand muscle using electromyography (MP160, BIOPAC). The AMT was defined as the lowest stimulator output in percentage that elicits 5 out of 10 twitches of more than 200 μV peak-to-peak amplitude in the contralateral hand, while the participant maintains 20% of a finger-thumb contraction. Previous studies (Stokes et al., 2005, 2007) have shown a positive relationship between scalp-cortex distance and motor threshold. Thus, the AMT found by the previous procedure was adjusted for differences in scalp–cortex distance between brain regions according to a simple linear formula:

$$adjAMT = AMT + m(DSiteX - DM1)$$

Being the adjAMT the adjusted AMT, AMT the unadjusted AMT, DM1 the distance between the scalp and the motor cortex (M1), and DSiteX the distance between the scalp and the target brain region, either DLPFC, PPC or S1 in this study. Finally, the m is the average distance gradient relating AMT to distance which we defined as 3 based on Stokes et al. (2005).

S2: Data inclusion

As determined by the standard safety protocols used for MRI and TMS experiments, participants were included if they were 20 years of age, not pregnant, not suffering claustrophobia, had no back problems that would interfere with their ability to recline in the scanner, had no ferromagnetic metallic or electronic implants in their body, had no medical history of convulsions or syncope, nor a family history of epilepsy, and they were not using any medications that were contraindicated for scanning (Hallett, Rossini & Pascual-Leone, 2009). Participants were recruited with normal or corrected-to-normal visual acuity, and without any neurological or psychiatric history.

S3: Data exclusion

Participants' objective performance was adjusted by using adaptive staircase to ensure their accuracy was approximately 75%, at each session of the experiment (Kaernbach, 1991). Accuracy levels below 60% would indicate that, for that block, the participant's performance is at chance level and would, therefore, need to be excluded from further analyses. If any participants had unstable session data sets (across 3 blocks), meaning type1 or type2 false alarm rate below FA <5% or hit rate above 95%, was excluded from analysis.

Treatment of outlier data: Following a recent cTBS meta-analytic study estimating a rather small effect size over the DLPFC (Lowe et al., 2018; which we approximate to a *Cohen's* d_z of 0.38) implying that extreme outlier values might shadow our parametric mean comparisons, we run the Grubb's test (Maximum Normed Residual test) as a rule to determine if any minimum or maximum value should be regarded as outlier and thus excluded from that particular statistical test.